Metabarcoding and Metagenomics 2: e21519
DOI: 10.3897/mbmg.2.21519

Research Article

Geographic distance and mountain ranges structure freshwater protist communities on a European scale

Jens Boenigk[1,2], **Sabina Wodniok**[1], **Christina Bock**[1], **Daniela Beisser**[1], **Christopher Hempel**[1], **Lars Grossmann**[1], **Anja Lange**[3], **Manfred Jensen**[1]

1 *University of Duisburg-Essen, Department of Biodiversity, Essen, Germany.*
2 *University of Duisburg-Essen, Centre for Water and Environmental Research, Essen, Germany.*
3 *University of Duisburg-Essen, Research Group Bioinformatics, Essen, Germany.*

Corresponding author: *Jens Boenigk* (jens.boenigk@uni-due.de)

Academic editor: *Agnès Bouchez* | **Received** 9 October 2017 | **Accepted** 14 December 2017 | **Published** 2 January 2018

Abstract

Protists influence ecosystems by modulating microbial population size, diversity, metabolic outputs and gene flow. In this study we used eukaryotic ribosomal amplicon diversity from 218 European freshwater lakes sampled in August 2012 to assess the effect of mountain ranges as biogeographic barriers on spatial patterns and microbial community structure in European freshwaters. The diversity of microbial communities as reflected by amplicon clusters suggested that the eukaryotic microbial inventory of lakes was well-sampled at the European and at the local scale. Our pan-European diversity analysis indicated that biodiversity and richness of high mountain lakes differed from that of lowland lakes. Further, the taxon inventory of high-mountain lakes strongly contributed to beta-diversity despite a low taxon inventory. Even though ecological factors, in general, strongly affect protist community pattern, we show that geographic distance and geographic barriers significantly contribute to community composition particularly for high mountain regions which presumably act as biogeographic islands. However, community composition in lowland lakes was also affected by geographic distance but less pronounced as in high mountain regions. In consequence protist populations are locally structured into distinct biogeographic provinces and community analyses revealed biogeographic patterns also for lowland lakes whereby European mountain ranges act as dispersal barriers in particular for short to intermediate distances whereas the effect of mountain ranges levels off on larger scale.

Key Words

protist, molecular diversity, biogeography, microbial ecology, algae

Introduction

Protists are not only a very diverse group of organisms but also quantitatively and qualitatively important components of all ecosystems. Their key ecological role has become a paradigm in microbial ecology based on the concept of the microbial loop (Azam et al. 1983, Boenigk and Arndt 2002, Pomeroy 1974): Prokaryotes essentially drive freshwater biogeochemical processes, whereas macroorganisms structure aquatic food webs and protists drive aquatic primary production and link algal and bacterial production to higher trophic levels. Autotrophic protists (algae) are the dominant primary producers in aquatic ecosystems and contribute roughly 50 % to global primary production, i.e. they are of similar importance as land plants. Heterotrophic protists, on the other hand, are the primary agents of the top-down control of bacteria and primary producers and are thus key organisms in determining the fate and transport of organic matter in nature (Boenigk and Arndt 2002, Caron 2001, Foissner 1987, Sherr et al. 2007) .

Increasing our knowledge of biodiversity distribution patterns at any level of the tree of life is of paramount ecological and evolutionary significance (Campo et al.

Copyright *Jens Boenigk et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

2016). This is of particular importance as distribution and diversity patterns are generally linked to ecosystem functioning and stability (Cardinale et al. 2012). In the past, protist distributrion has largely been investigated with a focus on ecological factors. Besides climate, the pH and the concentration of ions and nutrients are particularly important structuring factors on the local scale (Azovsky et al. 2016, Triadó-Margarit and Casamayor 2012, Wu et al. 2009). A large-scale plankton analysis demonstrated temperature to be a major driver of community composition in the marine (de Vargas et al. 2015) whereas nutrient concentrations and pH seem to be more important in freshwater communities (Tolotti et al. 2003, Tolotti et al. 2006, Triadó-Margarit and Casamayor 2012). Further, and particular along alpine gradients, elevation and environmental parameters co-varying with elevation such as UV-intensity also affect protist taxa and potentially shape communities (Sommaruga 2001, Sonntag et al. 2010).

In contrast to ecological factors the biogeographic distribution patterns of protists and the significance of historical factors potentially structuring their distribution came only recently into focus. In contrast, for macroorganisms biogeography and the factors driving distribution patterns of life on Earth have attracted scientists already for centuries. The realization that different geographic regions, despite similar ecological conditions, may be inhabited by different organisms goes back at least to Buffon (1707-1788) (Browne 1983) and became prominent as the first principle of biogeography. The importance of geographic barriers has been accepted since Humboldt's work on plant biogeography in 1807 (Humboldt and Bonpland 1807).

While biogeographic theory has developed into a strong conceptual framework for understanding the distribution patterns of animals and plants, its applicability to microbes has remained controversial. There have been disputes on the biogeographic distribution of microbial organisms since at least the early 19th century (Sprengel and de Candolle 1819). During the past few centuries, a uniform global dispersal has been proposed for microorganisms, culminating in the claim of Baas-Becking and Beijerink that "everything is everywhere, but the habitat selects" (Bass and Boenigk 2011, Wit and Bouvier 2006). Only recently has increasing evidence demonstrated that protists have a biogeography, and the underlying assumption of a worldwide dispersal of protists has been progressively replaced by the acceptance of the endemic distribution of at least some taxa (Bass and Boenigk 2011, Foissner 2006, Livermore and Jones 2015). The research focus has consequently recently turned to the underlying patterns of protist distribution while dispersal limitation is now generally accepted (Bik et al. 2011, Grossmann et al. 2016a, Katz et al. 2005, Fernández et al. 2017).

Most, if not all, microbial ecologists meanwhile agree that at least some single-celled organisms have limited distribution patterns (Filker et al. 2016, Kammerlander et al. 2015, van der Gast 2014, Wu et al. 2009) and that on a larger scale, historical and geographic factors may also affect protist distribution and dispersal (Filker et al. 2016, Izaguirre et al. 2015). However, despite indications of large-scale variation between protist communities, geographical factors are considered to be of minor importance at short to intermediate distances (Beisner et al. 2006, Filker et al. 2016). If this turns out to be true, protist distribution patterns would clearly deviate from those of macroorgansims, in particular in Europe due to the recent geological past of this continent: In the quarternary, Europe underwent several cycles of glaciations which strongly structured biological communities – at least of macroorganisms. Hewitt (2000) specifically stated that, "the present genetic structure of populations, species and communities has been mainly formed by Quaternary ice ages". In Europe, these patterns comprise centers of endemism, specifically in high mountain areas (Albach et al. 2006, Schmitt 2009), as well as genetically distinct lineages, mainly in the lowlands derived from different refugia (Santucci et al. 1998, Hewitt 2004, Babik et al. 2004). Europe is the probably best investigated area for biogeographic patterns (Hewitt 2000, Schmitt 2009) and therefore an ideal test case for further testing the importance of historic factors in structuring the biogeography of protists. The European orogenes – the Alps in the center of Europe and further orogenes towards the west (Pyrenees, Massif central), to the south (Apennine), to the east (High Tatra, Carpathian Mountains) and to the north (Scandinavian orogenes) constitute a unique area for glacial distribution barriers. Since the maximum of the last glaciation some 18,000 years ago (Litt et al. 2007, Habbe 2007) cold-adapted species took refuge in high mountain areas resulting in considerable endemism (Hewitt 2000, Schmitt 2009), whereas warm adapted species retracted to Southern Europe and re-colonised central Europe along several routes around the orogenes, yielding genetically differentiated populations in the lowlands.

Here we focus on the importance of the recent geological history to the distribution patterns of protists in lakes. For protists, geographic gradients are generally known (e.g. Casteleyn et al. 2010) but presumably weaker in freshwaters than in marine systems (Hillebrand 2004) or in soils Grossmann et al. (2016a). Nevertheless, even though geographic gradients are assumed to be weak in protists (Gimmler et al. 2016) historical factors presumably affect protist distribution as their distribution patterns in freshwater violate the assumption of unlimited dispersal (Grossmann et al. 2016a). In particular, we hypothesize that European mountain ranges act as dispersal barriers for protists. It is unknown to what extent diversity patterns as known for macroorganisms also apply to the less familiar and less readily observable protists inhabiting these same ecosystems (Mahé et al. 2014). To evaluate diversity patterns at the microbial scale for protists, we conducted a DNA metabarcoding study based on a large-scale comparison including 218 lakes across Europe.

Material and methods

Site selection and field sampling

We sampled 218 European freshwater lakes and ponds from sites in Norway, Sweden, Germany, Poland, the Czech Republic, Slovakia, Hungary, Romania, Austria, Italy, France, Spain and Switzerland. Site selection focused on the European orogens, specifically the Alps, the Pyrenees, the Apennine, the High Tatras, the southern Scandinavian mountains and the connecting flatlands (Suppl. materials 1, 2). We selected natural lakes (and reservoirs) which were typical for the respective area. Lakes which were known to be extremely deviate such as e.g. acidic mining lakes were excluded. Plankton samples were taken at a depth between 0.2 and 0.8 m near the shoreline and immediately processed after sampling. All samples were taken in August 2012. Temperature, pH and conductivity were measured directly in the field using a portable instrument. Climatic data including photosynthetically active radiation (PAR) and UV irradiation were partly measured on site and partly taken from databases of the Deutsche Wetterdienst and associated weather services. For molecular analyses, water samples were filtered onto 0.2 µm nucleopore filters (47 mm diameter) until the filters clogged, i.e. several hundred milliliter up to 600 ml water per filter. Subsequently, filters were air dried and then immediately frozen in liquid nitrogen (Cryoshippers). The filters were stored at -80 °C in the laboratory until further processing.

DNA extraction, PCR and sequencing

Genomic DNA was extracted using the my-Budget DNA Mini Kit (Bio-Budget Technologies GmbH) following the protocol of the supplier with the following modifications: Filters were homogenized in 800 µl Lysis Buffer TLS within lysing Matrix E tubes (MP Biomedicals) using the FastPrep instrument (MP Biomedicals). Homogenization was run three times for 45 seconds each at a speed setting of 6 m/s and then incubated for 15 min at 55 °C. The next steps followed the standard protocol supplied by Bio-Budget Technologies GmbH. The quality and quantity of the DNA was checked using a Thermo Scientific NanoDrop® ND-2000 UV-Vis spectrophotometer (Thermo Fisher Scientifics). PCR amplifications targeted the SSU V9 region and ITS1. Briefly, in order to cover a broad taxonomic spectrum, two primers with different wobble positions were combined in a ratio of 10% : 90%: 5'-GCTGCGCCCTTCATCGKTG-3' (ITS2_Dino; 10%) and 5'-GCTGCGTTCTTCATCGWTR-3' (ITS2_broad; 90%). We used the Amplicon-Duo pipeline (Lange et al. 2015) to separate true biological variation from sequence artefacts. For each sample, two technical replicates of the extracted DNA were independently amplified using primers with different sample identifiers [see Lange et al. 2015 for details]. The concentrations of the PCR reaction were as follows: 1 µl of DNA template (depending on the concentration, different dilutions of 1:1, 1:10, and 1:100 were used) in 25 µl PCR reaction with 0.4 units of Phusion Taq (Biozym), 0.25 µM primers, 0.4 mM dNTPs and 1 x Phusion buffer (Biozym). The PCR-cycling conditions included an initial denaturation step at 98 °C for 3 min followed by 35 cycles each including a denaturation step at 98 °C for 30 s, annealing step at 52 °C for 75 s, and an elongation step at 72 °C for 60 s. At the end the PCR was completed by a final extension step of at 72 °C for 10 min. Equimolar subsamples were pooled and commercially sequenced using paired-end HiSeq 2500 sequencing, applying 2x 300 bp reads using the "rapid run" mode on the Illumina platform of a sequencing provider (Fasteris, Geneva, CH).

Sequence filtering

Adapter and quality trimming was performed by the sequencing company. Samples were demultiplexed by the sequencing company (Fasteris) using MID sequences. Base quality of the sequence reads was checked using the FastQC software (Andrews 2015). A split-sample filtering protocol for Illumina amplicon sequencing was used as described in Lange et al. (Lange et al. 2015). In brief, the raw sequences were quality filtered to remove reads with an average Phred quality score below 25 using PRINSEQ-lite [v0.20.4, (Schmieder and Edwards 2011)]. Additionally, all reads with at least one base with a Phred quality score below 15 were removed provided that the sequences contained Ns after base calling. The paired-end reads were assembled and quality filtered with the tool PANDASeq [v2.10, (Masella et al. 2012)]. Reads with uncalled bases, an assembly quality score below 0.9, a read overlap below 20, or a base with a recalculated Phread-score below 1 were discarded. All reads were dereplicated, whereupon chimeras were identified using UCHIME [usearch v7.0.1090, (Edgar et al. 2011)] with default parameters. Finally, sequences that were not found in both sample branches using AmpliconDuo (Lange et al. 2015) were discarded.

Statistical analysis

Reads remaining after the AmpliconDuo filter were clustered via the software SWARM (version 2.1.9; Mahé et al. 2014), then clustered by identical V9 sequences (150 basepairs, ident = 100 %; R-Script "V9_Clust.R" by (Jensen 2017) and subsequently taxonomically assigned by searching the NCBI database using BLASTn. The taxonomic resolution therefore roughly corresponded to a 100% SSU V9 sequence identity. All reads assigned to Metazoa and Viridiplantae as well as higher fungi were excluded from further analyses, as protists are the target organisms of this study. The remaining read counts ranged between 19,267 (Lake Estany de Trebens) and 2,330,071 (Lake Millstätter See) per sample. OTU richness and abundance were computed from not transformed (raw) data; community structure indices

were also calculated based on the raw read abundance of OTUs. Rarefaction curves were plotted by the function "rarecurve" of the R-package vegan (version 2.4-1; Oksanen et al. 2017). In order to achieve a balanced dissimilarity matrix of all sites (accounting for differences in sequencing depth between the sites, etc.), we first applied Hellinger normalization. For comparison of community structures of the different sites we subsequently used the percentage difference dissimilarity matrix (Bray). We finally checked that the resulting dissimilarity matrix was euclidic (Dray et al. 2017). The resulting dissimilarity matrix was used for all further analyses. We calculated the local contribution of every site to beta-diversity (LCBD) (Dray and Dufour 2007) using again the percentage difference dissimilarity of the Hellinger-transformed data (see Suppl. material 3 for the Bray distance matrix). For further insights into the processes determining high betadiversity we also analysed the relative contribution of replacement of OTUs between lakes and abundance changes using the function beta.div.comp of the R-package adespatial (Legendre 2014b). By the parameters coef="S", quant=T this analysis was again based on Hellinger transformed (euclidic) Bray-distance data. Geographic groups of community compositions were calculated by geo-constrained hierarchical clustering analysis (Legendre 2014) resulting in 14 geographic clusters. The R-package tripack (Renka et al. 2016) was used for construction of the geographic neighbour matrix based on Delauney triangulation. The effect of geographic distance was investigated separately for lowland and high mountain lakes (see supplementary meterial 4). We calculated the distances [km] of all site pairs as well as the community distances of all site pairs (submatrices of the community dissimilarity matrix) and analyzed the effect of geographic distance by linear regression after conversion of geographical latitude and longitude data into distances [km] (Chambers 2013). The effect of mountain ranges (in particular of the Alps and the Pyrenees) as geographic barriers was also analyzed by linear regression based on splitting the sample set into three groups: an alpine group including the Alps, the Pyrenees and the High Tatras, a northern group (cis) and a southern group (trans). The mountain group was identified by altitudes plus constrained clustering (see above). By exclusion of the alpine group we could compare subsamples, i.e. all subsamples "cis-cis" (northern), all subsamples "trans-trans" (southern) and all cis-trans comparisons. Mean differences of the cis-cis, trans-trans, cis-trans groups were tested by t-tests. All figures were prepared with R and Adobe Illustrator.

Data resources

All sequences are deposited at Genbank (accession number PRJNA414052).

Results

Protist freshwater communities show high levels of diversity

Sequencing of the European freshwater samples resulted in 323,430,198 cleaned V9 forward and reverse read pairs. We applied the AmpliconDuo filter (Lange et al. 2015) for separating true biological diversity from PCR and sequencing artefacts. This novel cleaning step based on technical replication effectively discarded sequence artefacts, resulting in a total number of 100,516,809 reads. Operational taxonomic units were constructed with the clustering algorithm SWARM, resulting in 74,713 distinct unique V9 sequence SWARMs (i.e., clusters). 64,969 V9 SWARMs (95,592,035 reads) of these unique sequences were assigned to protists. The remaining 13 % of V9 SWARMs (4.9 % of reads) which were excluded

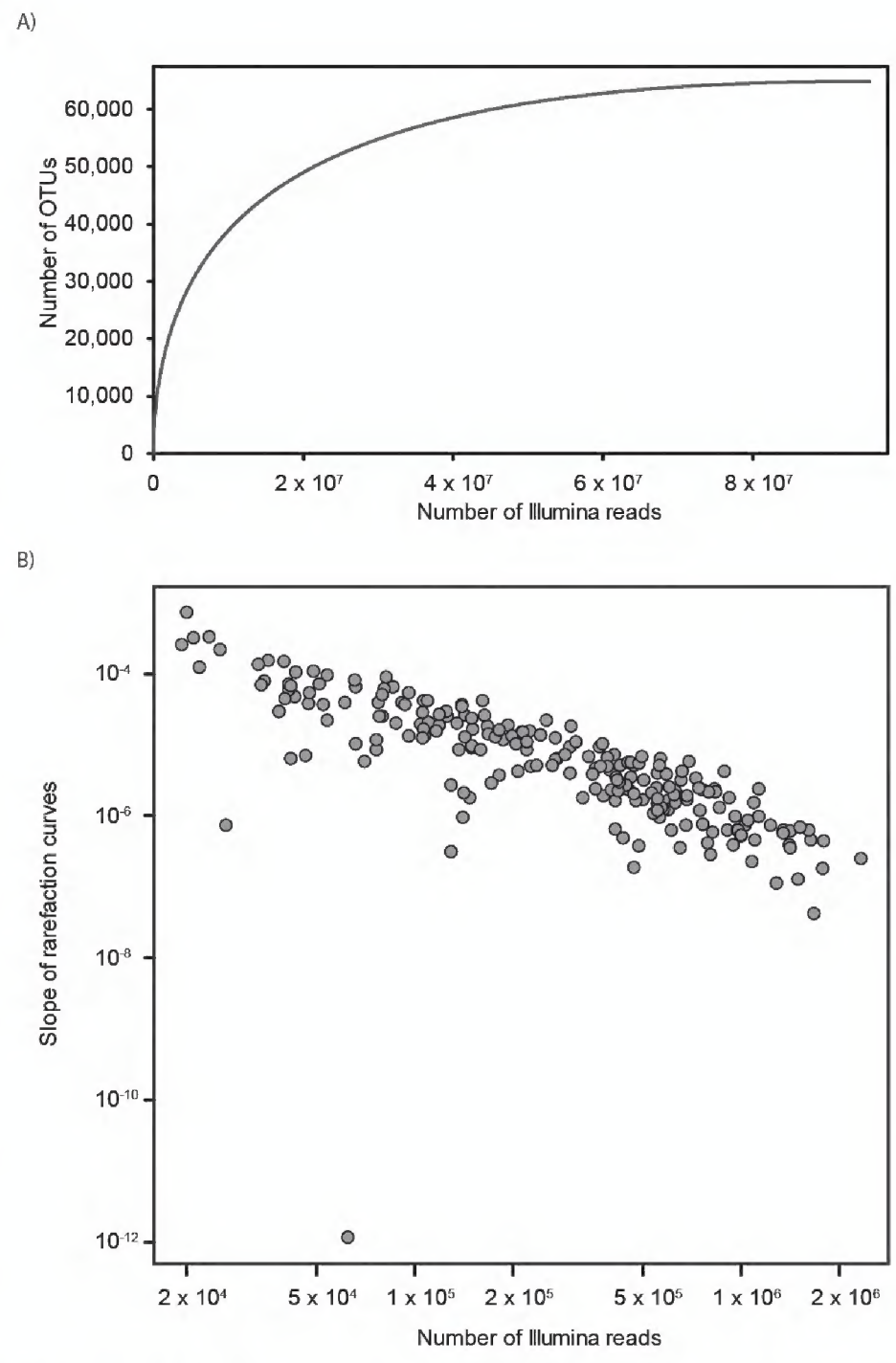

Figure 1. Planktonic protist ribosomal diversity. (A) V9 rDNA OTUs rarefaction curve. (B) Saturation slope versus number of V9 rDNA reads for all of the 218 samples analyzed herein. A slope of 10^{-6} indicates that one novel sequence read can be recovered if one million new reads are sequenced. Overall, the saturation values indicate that the extensive sampling effort uncovered the majority of eukaryotic ribosomal diversity both at the European and the local level.

from further analysis were affiliated with Metazoa (1,102 V9 SWARMs; 638,687 reads), Embryophyta (599 V9 SWARMs; 291,307 reads), the fungal lineages Basidiomycota, Ascomycota and Glomeromycota (5,110 V9 SWARMs; 3,856,310 reads) or not affiliated with the domain Eukarya (2,933 V9 SWARMs; 138,470 reads).

The applied sequencing depth was sufficient to approach saturation of eukaryotic taxon richness both at the local and European scale (Fig. 1) and resulted in 2 x 10^4 to 2 x 10^6 cleaned reads per sample (average 438,495 reads per site). At the European scale, the rarefaction slope dropped below 10^{-4} above 63,693 OTUs corresponding to 67,000,000 reads. Towards the end of the rarefaction curve the slope was 4.44 x 10^{-8}, correspondingly one new OTU can be expected for 22.5 million new reads. Using the same saturation criterion, i.e. a rarefaction slope of 10^{-4}, local richness approached saturation between several hundred and 4,000 OTUs, corresponding to approximately 0.5 – 6 % of the European scale richness. At the level of taxonomic groups, ciliates (34.6 %) followed by dinoflagellates (12.7 %), green algae (12.0 %), cryptophytes including katablepharids (8.1 %), chrysophytes (6.6 %) and diatoms (4.2 %) contributed most to richness (Fig. 2). The individual OTUs contributing most strongly to overall eukaryotic reads were affiliated with an uncultured katablepharid, the green alga *Desmodesmus insignis*, and several ciliates.

High mountain lakes differ from lowland lakes in taxon inventory and act as biogeographic islands

The OTU pool showed a good fit to the log-normal distribution (Fig. 3a). Variation of V9 OTU richness between distinct lakes was high, with a mean of 438,496 ± 423,896 V9 reads per lake. Community overlap between lakes was generally low, with only a few widespread taxa (Fig. 3b, Suppl. material 1). Taxon inventory differed with elevation, and high mountain lakes had a particularly low taxon overlap with lowland lakes. High mountain lake communities thus distinctively differed from their surroundings. Accordingly, the strongest community dissimilarities between neighbouring lakes occurred along mountain ranges (Fig. 4, Suppl. material 2). Furthermore, different high mountain areas such as the Pyrenees, the Alps, and the High Tatras decisively differed in taxon inventory from each other. High mountain areas are thus biogeographic islands to protist distribution.

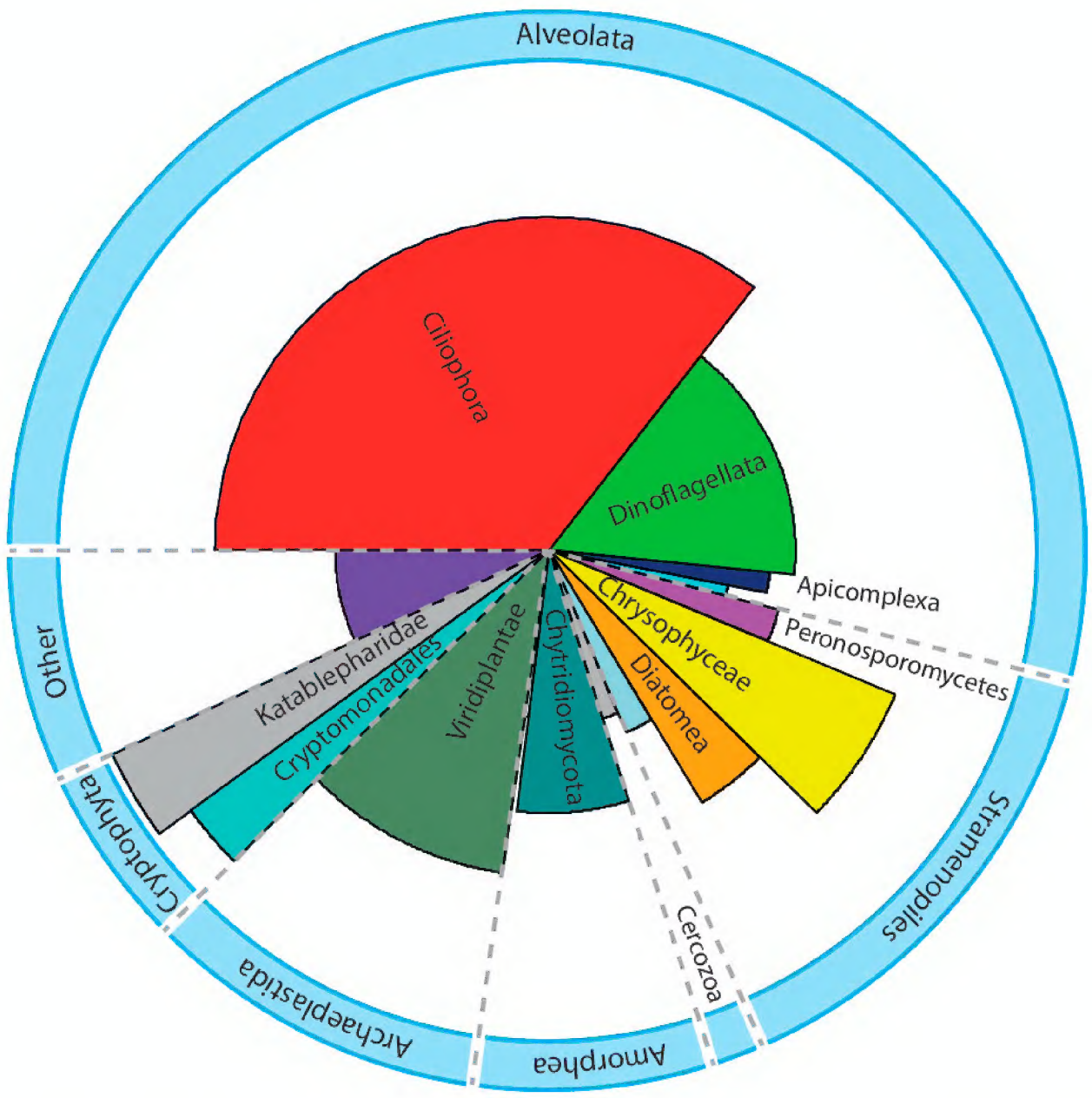

Figure 2. Phylogenetic affiliation of reads and OTUs to eukaryotic taxonomic groups. The area of the portion of the pie chart represents the number of reads, whereas the angle represents the relative contribution to OTU richness. Alveolates and in particular Ciliophora and Dinoflagellata contributed nearly 50 % to global OTU richness. Cryptophyta, Viridiplantae (without Embryophyta), as well as Chrysophyceae and Diatomea accounted for the majority of the remaining OTUs. Unlabeled pieces within the eukaryotic supergroups comprise all other taxa affiliated with that group.

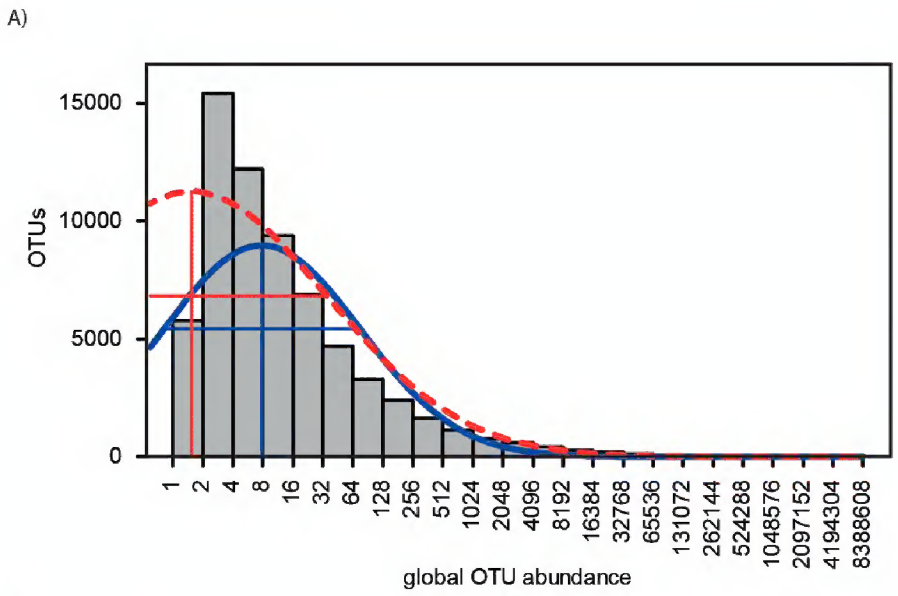

Figure 3. Abundance distribution and ubiquity of OTUs. A) Global OTU abundance distribution and fit to the Preston log-normal model. Quasi-Poisson fit to octaves (red curve) and maximized likelihood to log2 abundances (blue curve). Most OTUs were represented by 3 to 32 reads. Approximations fit to the Preston log-normal model. The maximized likelihood to log2 abundances (blue curve) fit was superior and subsequently used to calculate the Preston veil, which infers the number of OTUs that we missed during our sampling. Preston veil was 8,111, confirming that we captured most of the diversity as also indicated by the rarefaction slopes. Preston vail and rarefaction slopes confirm that holistic and general patterns of eukaryotic freshwater diversity can be extracted from our data. The unit of the y-axis is the number of OTUs B) Distribution of OTUs among European lakes. Few OTUs were ubiquitous (occurring in the majority of lakes) whereas most OTUs were specific to one or only a few lakes. Unit of y-axis as in a).

Beyond a shift in taxon inventory with elevation, richness pronouncedly decrease around 1400m (Suppl. material 4) and V9 OTU richness in high mountain lakes above 1,400m (456 ± 439) was generally lower than in low elevation lakes below 500m (843 ± 653). Most V9 OTUs were specific to a few sites, and thus lakes with a high proportion of OTUs with restricted distribution considerably contributed to beta-diversity. Contribution

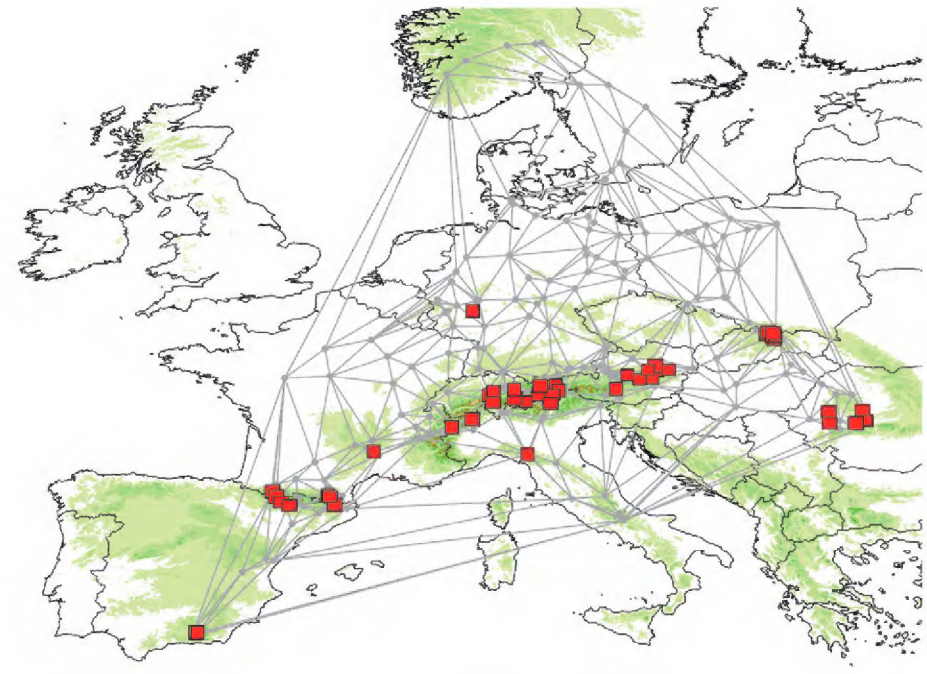

Figure 4. Delaunay triangulation plot based on the investigated lakes. The red squares indicate triangles of high dissimilarity (i.e. larger than mean + one standard deviation) between the three corner sampling sites. Community dissimilarities were calculated based on wheighted distances using the distance function x=((1-distance)/maximal triangle leg). The areas of high dissimilarity occur nearly exclusively along mountain ranges, in particular along the Alpes, the Pyrenees, the High Tatras, the Carpathian mountains and the Sierra Nevada.

to beta-diversity was particularly high in high mountain regions (Fig. 5a), indicating a higher degree of habitat isolation due to either ecological factors, i.e. habitat heterogeneity, or geographic factors, i.e. isolation due to geographic barriers. The high beta-diversity data and the high values within the distance matrix indicated that different lakes harbour primarily different OTUs. The analysis of beta-diversity for all lakes indicated that 75.5 % of beta diversity was due to replacement and 24.5 % of beta diversity was due to abundance differences. When the analysis was restricted to lowland lakes of the northern area (cis) the contribution of replacement was 73.3 % and for the lowland lakes of the southern area (trans) to 79.3 %. For lakes above 1400 m beta-diversity was increased and 79.7 % of this beta diversity was due to replacement but only 20.3 % due to abundance differences.

When restricting the analysis to presence/absence data most of the beta-diversity was due to replacement and only a minor part due to loss of species (nestedness), i.e. 4,0 % for all lakes. However, the nestedness was somewhat higher within the northern low elevation (cis) group, i.e. 5.4%, possibly (even though hypothetical) to incomplete recolonization since the last glaciation events whereas in high elevation lakes which possibly act as refuge the nestedness is very low, i.e. 2.1 %.

Geographic distance shapes protist distribution patterns and mountain ranges act as barriers to protist dispersal

On the European scale environmental factors slighty co-varied with distance for lowland lakes but did not systematically change with distance for high mountain lakes (Fig. 6A). In contrast, community similarity generally

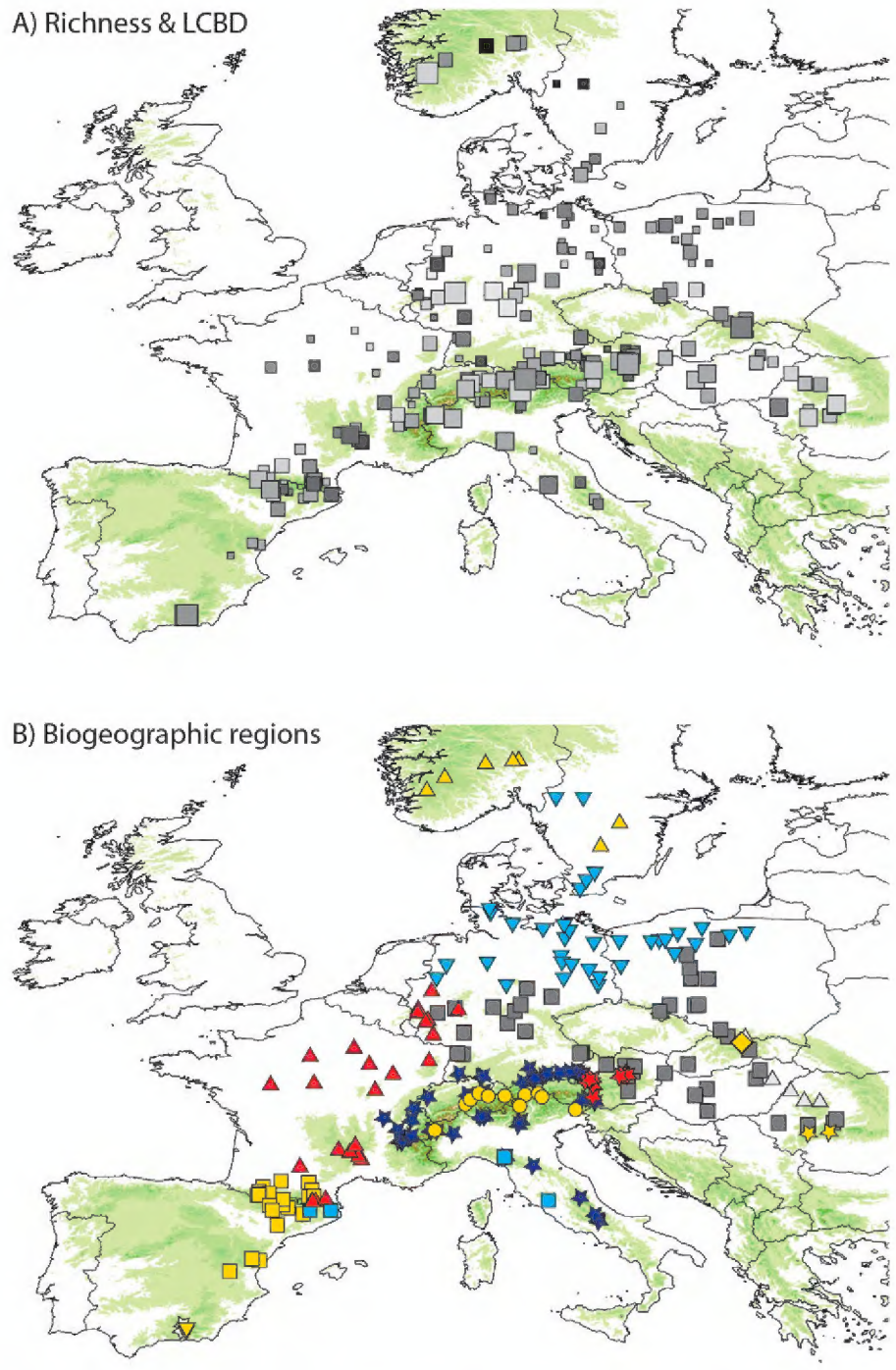

Figure 5. OTU distribution and protist biogeographic regions in Europe. A) Richness and contribution to beta diversity of individual lakes. Richness is illustrated by grey scale of the symbols: dark shades indicate a high richness whereas light shades indicate a low richness. The size of the symbols indicates the relative contribution to beta diversity, with large symbols indicating a high contribution. Richness of individual lakes varied between 15 and 4092. Lakes in high mountain regions tended to contribute at an above-average level to beta diversity. B) Protist biogeographic regions as calculated by geo-constrained hierarchical clustering. High mountain ranges appear as biogeographic islands; the geographic regions largely comprising high mountain protist communities are depicted in yellow. Furthermore, the lowland areas are biogeographically structured which is mainly due to an east- west-gradient. Even though mountain ranges affect protist community similarity between sites, they do not necessarily separate biogeographic regions.

decreased with increasing distance both for highland and lowland sites (Fig. 6B). A co-inertia analysis using the R-package ADE4 of the normalized community structures (Hellinger-Bray transformed) versus the scaled environmental variables pH, temperature, and log conductivity indicated no strong correspondence between community composition and environmental factors (RV value of 0.22) thus corroborating a contribution of geographic distance and historic factors to patterns community composition.

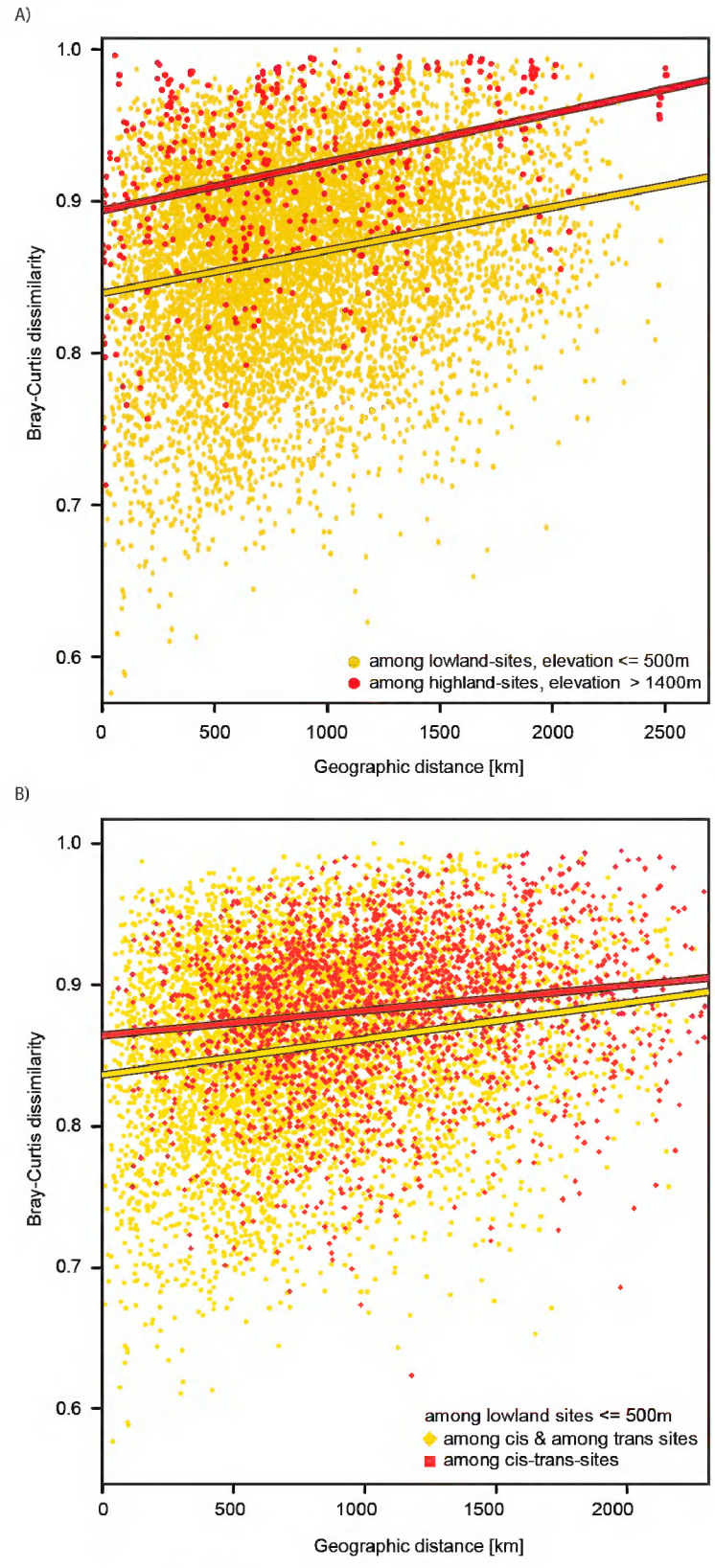

Figure 6. Effect of geographic distance on community dissimilarity. A) Habitat dissimilarity for high mountain lakes (above 1,400 m) and for lowland lakes (below 500 m). The latter were either separated (cis-trans pairs) or not separated by mountain ranges (cis-cis pairs, i.e. north of the Alps, Pyrenees and High Tatras and trans-trans pairs, i.e. south of the Alps, Pyrenees and High Tatras). Regression lines are shown for highland (blue), lowland cis-cis and trans-trans pairs (yellow), and lowland pairs separated by mountain ranges (orange). Habitat dissimilarity increased with distance for lowland lakes but did not change for high mountain lakes. B) Bray-Curtis community dissimilarity for the same groups as in A. Community dissimilarity in high mountain regions increased with distance even though habitat dissimilarity did not change with distance. For lowland lakes both, habitat characteristics and community composition, changed with distance. However, the effect of distance on habitat dissimilarity was independent on whether the lakes were separated by mountain ranges or not. In contrast, community dissimilarity was higher at low and intermediate distance for lakes separated by mountain ranges. Accordingly the slope of the regression line for cis-trans pairs was significantly lower as for cis- cis and trans-trans pairs ($p = 0.0097$). The effect of geographic barriers levels off for longer distances above approximately 1,500 km.

Beyond the general effect of geographic distance, geographic barriers such as mountain ranges contributed significantly to community dissimilarity between lowland sites on a European scale. Mean community similarity was higher for lakes within a given lowland area as compared to community similarities of lakes between areas which were separated by a mountain range (Fig. 6B). For instance, separate analyses of sites north of the Alps and south of the Alps resulted in comparatively high community similarities over a short distance, and a pronounced decrease of community similarity with geographic distance. In contrast, community similarity was generally low for habitat pairs across a mountain range already for short distances and geographic distance had a minor effect on community similarity in the latter case. Thus biogeography, in particular geographic distance and geographic barriers, shapes protist distribution pattern. Interestingly, distant lakes still share a fraction of their taxon inventory, clearly demonstrating that geographic distance and geographic structures such as mountains are not absolute barriers to protist dispersal. The effects of geographic distance and geographic barriers are, however, not strictly additive. In contrast, both factors add to community dissimilarity up to a certain level of background community similarity shaped by environmental variables.

Protist diversity in European lakes is structured into biogeographic regions

Despite a generally low community similarity between different lakes, groups of lakes sharing a higher level of community similarity can be identified. Lakes sharing a higher degree of community similarity form clusters reflecting biogeographic provinces on a European scale (Fig. 5b). Notably mountain ranges are borders to protist biogeographic provinces as for animals and plants as well. This pattern was consistent in particular for the Alps, the Pyrenees and the High Tatras. The taxon inventory of high mountain regions strongly deviates from that of lowland regions, and lowland areas are separated from each other based on their taxon inventory into distinct provinces. These provinces do not correspond to physicochemical patterns or to the distribution of soil types or geology in Europe, suggesting that factors related to geographic distance and dispersal are key reasons for this separation. It must be noted that towards the edges of the investigation area the delimitation of such biogeographic regions become more uncertain due to the shortage of neighboring sampling sites or sites in between. For instance, whether the lakes indicated by the blue squares in eastern Spain and northern Italy represent a connected biogeographic province or rather two separate areas which share a high similarity by chance remains uncertain.

The decreasing effect of mountain ranges in separating protist communities with distance between sites indicates that geographic barriers act only as relative dispersal barriers for protists; they do not inhibit dispersal, but rather slow it down and thereby largely act on short to intermediate distances. For larger distances, the separating effect of geographic barriers adding to that of geographic distance alone largely vanishes, as demonstrated by cis-cis and trans-trans (i.e. within a distinct lowland area) versus cis-trans (i.e. lowland sites separated by a mountain range) comparisons of habitat dissimilarities (Fig. 6B).

Discussion

The temporal and spatial pattern of eukaryotic microbial diversity is underexplored and has only recently come into focus (Green et al. 2004, Lara et al. 2011, Livermore and Jones 2015, Nolte et al. 2010). Even though ecological factors are largely responsible for structuring protist communities, we found indications that community composition is possibly imprinted by deeper, older patterns reflecting the recent geological history and the geographic separation by mountain ranges in europe.

Eukaryotic microbial molecular diversity in European freshwaters

Protist diversity is undoubtedly tremendous (Bass 2004, Countway et al. 2005, Countway et al. 2007, de Vargas et al. 2015, Moreira and López-García 2002). Considerable research effort has been focused on marine diversity (Countway et al. 2005,Countway et al. 2007, Campo et al. 2016, Lovejoy et al. 2006, López-García et al. 2001, Massana and Pedrós-Alió 2008, Worden et al. 2006). Despite the global importance of freshwater resources and the heterogeneity of freshwater systems, their biological diversity has been comparatively neglected (Filker et al. 2016, Grossmann et al. 2016a, Kammerlander et al. 2015, Lepère et al. 2007, Simon et al. 2015, Taib et al. 2013, Triadó-Margarit and Casamayor 2012, Slapeta et al. 2005). Our study revealed a high diversity of microbial freshwater eukaryotes and a high level of regional separation. The majority of OTUs was affiliated with alveolates, indicating a high diversity within this group in particular. Within alveolates the Ciliophora showed the highest richness and comprised both, taxa with a restricted distribution and taxa with a broad distribution. In contrast, dinoflagellates and chytridiomycetes comprised a high number of OTUs specific to one or a few lakes indicating a high degree of specialization, whereas in Cryptomonadales and Katablepharidae most OTUs were generalist taxa occurring in many lakes.

Although comparatively smaller, freshwater systems are usually more heterogeneous than marine systems and offer a larger array and diversity of ecological niches (Simon et al. 2015). Freshwater communities are therefore expected to be much more differentiated between sites than marine communities (Barberán et al. 2011, Logares et al. 2009). In fact, we found a low taxon overlap between lakes and a correspondingly high contribution to beta-diversity of individual habitats. In accordance with our findings, a large genetic divergence between distinct

lake microbial communities has been reported in recent deep-sequencing campaigns, in particular for the rare biosphere (Filker et al. 2016, Grossmann et al. 2016b, Triadó-Margarit and Casamayor 2012). In particular, distribution patterns of rare taxa seem to vary widely and to deviate from the rather uniform patterns in abundant protists (Grossmann et al. 2016a, Schiaffino et al. 2016). This confirms recent claims that distribution patterns are almost certainly differential, with some taxa being ubiquitous whereas others are rather endemic, as proposed for instance by the moderate endemicity model.

Our survey aimed to cover all eukaryotic lineages, but universal primers may miss some relevant taxonomic groups (Massana et al. 2015). Distribution patterns are further obscured by partly inappropriate species delimitation in protists (Mann and Vanormelingen 2013, Scoble and Cavalier-Smith 2014). Appropriate species delimitation (and thus species distributions) are even more nebulous for the vast number of (so far) uncultivatable protist lineages (Massana et al. 2013). Assertions on the absolute species numbers and species inventory are therefore imprecise and subject to numerous assumptions. It is therefore important to keep in mind that some taxonomic groups are underrepresented and may exhibit patterns deviating from the general trend. Keeping this in mind, the community composition corresponded to those known from other studies (e.g. Lange et al. 2015, Kammerlander et al. 2015, Filker et al. 2016).

Our analysis based on 218 European lakes yielded in 74,713 distinct unique V9 sequence SWARMs. Rarefaction analysis as well as Preston vein indicated that we captured most of the diversity. However, molecular diversity inventories based on high throughput sequencing techniques must be interpreted with caution as high throughput sequencing platforms are error prone and in particular for sequences exclusively present in one or very few samples PCR artefacts and sequencing errors must be taken into account (Degnan and Ochman 2011, Lange et al. 2015). The AmpiconDuo filter largely removed such sequences by splitting the samples followed by two independent PCRs and sequencing steps for each sample. Only errors that occurred independently in both branches of the split sampling protocol may pass the filter. Even though it is highly unlikely that high numbers of artificial sequences pass the filter it is known that some artefacts such as formation of some chimeric sequences may, in part, occur systematically (Fonseca et al. 2012) and thus escape the filter. In fact, particularly some chimeric sequences pass the AmpliconDuo filter and not all of these sequences are recognized by chimera filters such as UCHIME (Lange et al. 2015). Therefore it is likely that some artificial sequences passed the filter steps, even though the share of sequence artefacts is probably lower as compared to past standard filtering algorithms. However, erroneous sequences should mostly be restricted to a single sample and thus not or hardly affect our analyses of community pattern across samples. The high number of OTUs restricted to a single lake may, nevertheless, be overestimated. It is likely, however, that due to the strict filtering steps the majority of these OTUs represent true biological diversity.

For some samples the number of OTUs was considerably lower as expected. This may be due to several factors: First, our filter strategy, including the AmpliconDuo filter and SWARM reduced the number of OTUs considerably. In particular, AmpliconDuo filtered out many presumably artificial sequences most of which had low to medium read numbers. Therefore the total number of OTUs should be expected to be somewhat smaller as compared to studies which do not apply this or a comparable filter (Lange et al. 2015). Second, the comparatively long amplicon (V9 + ITS1) and the long primer constructs also may miss some sequences, as it is generally known for long primer constructs and long amplicons (Huber et al. 2009). Third, some samples were strongly dominated by one or very few taxa due to algal blooms or strong dominance of a distinct bacterivore. For instance, Lake Pfordter See which was strongly dominated by a katablepharid yielded in as few as 15 OTUs. In such communities a lower taxon coverage was to be expected which may even have been self-inforcing as the dominance of one distinct sequence may cause a negative selection of rare sequences during PCR thereby further decreasing detection limits. Even though the total number of OTUs was certainly lower than the actual number of species present, the data structure as well as observations indicate that biological reasons, i.e. community structure, is presumably largely responsible for low OTU numbers in these lakes. In sum, even though in some lakes OTU number was surprisingly low, data structure and observations during sampling indicate that a strong dominance by few OTUs in these lakes is presumably largely real. Additional taxa, even though presumably extremely rare, are certainly to be expected. However, for the comparative analyses the applied Hellinger transformation balanced potentially deviating distributions and the analysis of distribution patterns is much less affected when based on defined units of biological diversity such as the herein-used V9 OTUs.

Contribution of biogeography and historical factors to the differentiation of freshwater communities

Ecological factors, in particular abiotic factors such as nutrient concentrations, temperature and light and UV intensity are well known to affect protist community composition (e.g Sommaruga 2001, Sonntag et al. 2010, Tolotti et al. 2003, Tolotti et al. 2006, Triadó-Margarit and Casamayor 2012). However, recent data provide increasing evidence that beyond the ecological factors also geographic distance and historical factors affect protist distribution at least on large scales (Filker et al. 2016, Schiaffino et al. 2016, Fernández et al. 2017). Europe is a well-investigated (if not the best-investigated) area for biogeographic patterns (Hewitt 2000, Schmitt 2009). The European mountain ranges – the Alps in the center of Europe and further mountain ranges towards the

west (Pyrenees, Massif Central), the south (Apennine), the east (High Tatras, Carpathian Mountains) and to the north (Scandinavian mountain ranges) – constitute a unique area for studying the effect of distribution barriers.

Based on the largest freshwater data set known to us, we demonstrate a significant contribution of geographical distance and historical factors to protist distribution patterns. We found spatial patterns of protist diversity with microbial community composition varying both with altitude in mountain ranges and with geographical distance. Seasonal taxon fluctuations influence protist community composition and may interfere with spatial analyses when samples originate from different seasons (Bock et al. 2014, Massana et al. 2015, Nolte et al. 2010). Considering that all samples in our campaign were taken in August 2012 and that seasonal effects and plankton succession dynamics is presumably lower in mid-summer lake plankton communities than in spring succession communities (Naselli-Flores et al. 2003), the observed divergence presumably reflects largely spatial differentiation of lake plankton communities rather than seasonal variation (Grossmann et al. 2016b). However, due to the large area included in the analysis a certain bias due to differential stages of the seasonal succession cannot be excluded.

Recently, de Vargas et al. 2015 have shown that the correlation between community similarity and geographic distance in marine ecosystems levels off for distances greater than 6,000 km. Our data confirm a similar relationship for limnic lowland systems which act, however, at shorter distances and is modified by local and regional geography. In particular, Bray-Curtis dissimilarity (supplementary material 3) increased with distance, but this correlation already levelled off for large distances and virtually vanished for sampling sites separated by geographic barriers. This indicates that protist freshwater communities in Europe are presumably largely driven not by geographic but by ecological factors.

For instance, we found a general trend of decreasing richness for high elevation lakes, i.e. high mountain lakes at elevations above approximately 1400 m showed a lower mean richness as compared to lakes at low elevations. This may be due to either the generally small size of mountain lakes as would be expected from species-area-relationships or indicate differential patterns of richness depending on elevation. As small lakes at low elevation can show a high richness we suspect that both, habitat size and elevation, or factors co-varying with elevation, may be repsonsible for the low richness.

Even though ecological factors are certainly of primary importance explaining for protist community composition, the contribution of geographic distance and historical factors (or geographic barriers) seems to have a stronger impact on protist freshwater communities than on marine communities (cf. de Vargas et al. 2015). Further, for high mountain ranges the effect of geographic distance is more pronounced as these regions are more isolated and may act as biogeographic islands at least to species with restricted dispersal capabilities. In this context it must be noted, that a considerable portion of protist distribution pattern is not explained by geograhic separation indicating that it is either restricted by other factors, i.e. presumably ecology, or not restricted by the geographic barriers, i.e. suggesting that the significance of geographic barriers may strongly differ between taxa. However, our data provide evidence that historical factors contribute to shaping protist distribution pattern (Fernández et al. 2016) beyond the presumably predominant effects of ecological factors.

According to the generally high beta diversity, our study revealed a high proportion of taxa exclusively found in only one or a few lakes. The very low contribtion of nestedness of the high elevation lakes and the somewhat higher contribution of nestedness to beta diversity in the northern lowland lakes may be interpreted as historical signals of the last glaciation and the post-glacial colonisation. However, as these differences are weak the interpretation as a historical signal remains hypothetical. How far postglacial resettlement, geographical barriers, other historical reasons, ecological parameters like nutrient supply or just biotic influences by other organismic groups dominate the exceptionality of the protist communities within lakes, remains to be shown. Further, it is likely that ecological and historical factors interact in shaping protist distribution pattern (Fernández et al. 2017). Irrespective of protist community differentiation due to ecological factors and geographic distance, we found a weak but nevertheless significant effect of mountain ranges as geographic barriers contributing to the separation of low elevation lakes (fig. 6). Future large scale analyses including cross-continental comparisons of such large datasets, in particular of lakes of a similar trophic status, will be a good test to confirm large scale patterns of geographic differentiation in protists and to differentiate regional from global patterns.

Conclusions

Mountain ranges have long been accepted as dispersal barriers for larger animals and plants, but their significance as dispersal barriers for microbes remained unclear. We demonstrate a differential impact of geographic distance for lowland and for highland lakes. Biodiversity and richness of high mountain lakes differ from that of lowland lakes. In accordance with the patterns for animals and plants (Nogués-Bravo et al. 2008, Romdal and Grytnes 2007), the mean protist richness was lower in high mountain lakes than in lowland lakes. Due to a high variation of the taxon inventory of individual lakes, this correlation only becomes apparent for high sampling numbers and therefore does not contradict more confined studies on smaller sample sets (Grossmann et al. 2016b). However, as a particularly high fraction of taxa in high-mountain lakes is presumably specific to only one or just a few lakes, the high-mountain lakes strongly contribute to beta-diversity despite the low taxon inventory.

Furthermore, mountain ranges act as islands of high mountain protist biodiversity and geographic distance accordingly clearly contributed to community dissimilarity. However, it must be noted that the overall contribution of geography to protist distribution pattern in lakes was low but, nevertheless, significant. Other factors such as abiotic and biotic factors presumably contribute to community composition much stronger as geography as it has been suggested in numerous studies (Beisner et al. 2006, Filker et al. 2016 and references therein). With respect to lowland lakes mountain ranges act as barriers to the dispersal of protist biodiversity, as already known for macroorganisms (Schmitt 2009). Thus, protist community dissimilarity is high even at short distances when geographic barriers separate the habitats. In this respect protist distribution patterns are similar to those of metazoans and land plants. However, the comparatively low effect of mountain ranges indicate that geographic and historical factors, even though certainly present, are less pronounced in protist communities as compared to plant and animal communities. Mountain ranges can block direct dispersal routes of microbes or at least of some microbial taxa, but a clear separation of biogeographic regions as observed for animals and plants is much less pronounced and levels off for lowland lakes on larger scales. Nevertheless, geographic structures affect protist distribution patterns on a European scale as reflected by distinct biogeographic regions.

Acknowledgements

We thank the German Research Foundation (projects BO 3245/19-1 and BO 3245/17-2) for financial support. We thank Joachim Stadel, Verena Stadel, Susann Chamrad, Steffen Jost, and Geoffrey Ongondo for support during the sampling campaign. We thank the Department Ecology, Biodiversity & Evolution of Animals at the University Salzburg, the Department Plant Ecophysiology at the University Konstanz, the Institut de Ciències del Mar in Barcelona, the Division of Clinical Physiology at the University of Debrecen, the University of Potsdam and the IGB Neuglobsow for supplying liquid nitrogen to the teams during the sampling campaign.

References

Albach DC, Schönswetter P, Tribsch A (2006) Comparative phylogeography of the *Veronica alpina* complex in Europe and North America. Molecular Ecology 15(11): 3269–3286. https://doi.org/10.1111/j.1365-294x.2006.02980.x

Andrews S (2015) A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed on: 2017-9-27]

Azam F, Fenchel T, Field J, Gray J, Meyer-Reil L, Thingstad F (1983) The ecological role of water-column microbes in the sea. Marine Ecology Progress Series 10: 257–263. https://doi.org/10.3354/meps010257

Azovsky A, Tikhonenkov D, Mazei Y (2016) An Estimation of the global diversity and distribution of the smallest eukaryotes: biogeography of marine benthic heterotrophic flagellates. Protist 167(5): 411–424. https://doi.org/10.1016/j.protis.2016.07.001

Babik W, Branicki W, Sandera M, Litvinchuk S, Borkin LJ, Irwin JT, Rafiński J (2004) Mitochondrial phylogeography of the moor frog, *Rana arvalis*. Molecular Ecology 13(6): 1469–1480. https://doi.org/10.1111/j.1365-294x.2004.02157.x

Barberán A, Fernández-Guerra A, Auguet JC, Galand PE, Casamayor EO (2011) Phylogenetic ecology of widespread uncultured clades of the Kingdom Euryarchaeota. Molecular Ecology 20(9): 1988–1996. https://doi.org/10.1111/j.1365-294x.2011.05057.x

Bass D (2004) Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of Cercozoa (Protozoa). International Journal of Systematic and Evolutionary Microbiology 54 (6): 2393–2404. https://doi.org/10.1099/ijs.0.63229-0

Bass D, Boenigk J (2011) Everything is everywhere: a twenty-first century de-/reconstruction with respect to protists. Biogeography of Microscopic Organisms. Cambridge University Press, Cambridge, 88–110. https://doi.org/10.1017/cbo9780511974878.007

Beisner BE, Peres-Neto PR, Lindström ES, Barnett A, Longhi ML (2006) The role of environmental and spatial processes in structuring lake communities from bacteria to fish. Ecology 87(12): 2985–2991. https://doi.org/10.1890/0012-9658(2006)87[2985:troeas]2.0.co;2

Bik HM, Sung W, de Ley P, Baldwin JG, Sharma J, Rocha-Olivares A, Thomas WK (2011) Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. Molecular Ecology 21(5): 1048–1059. https://doi.org/10.1111/j.1365-294x.2011.05297.x

Bock C, Medinger R, Jost S, Psenner R, Boenigk J (2014) Seasonal variation of planktonic chrysophytes with special focus on Dinobryon. Fottea 179–190. https://doi.org/10.5507/fot.2014.014

Boenigk J, Arndt H (2002) Bacterivory by heterotrophic flagellates: community structure and feeding strategies. Antonie van Leeuwenhoek 81(1–4): 465–480. https://doi.org/10.1023/A:1020509305868

Browne J (1983) The secular ark. Studies in the history of biogeography. Yale University Press 27(4): 452–453. https://doi.org/10.1017/s0025727300043611

Campo Jd, Guillou L, Hehenberger E, Logares R, López-García P, Massana R (2016) Ecological and evolutionary significance of novel protist lineages. European Journal of Protistology 55: 4–11. https://doi.org/10.1016/j.ejop.2016.02.002

Cardinale B, Duffy JE, Gonzalez A, Hooper D, Perrings C, Venail P, Narwani A, Mace G, Tilman D, Wardle D, Kinzig A, Daily G, Loreau M, Grace J, Larigauderie A, Srivastava D, Naeem S (2012) Biodiversity loss and its impact on humanity. Nature 486(7401): 59–67. https://doi.org/10.1038/nature11148

Caron D (2001) Protistan herbivory and bacterivory. Methods in Microbiology. https://doi.org/10.1016/s0580-9517(01)30050-8

Casteleyn G, Leliaert F, Backeljau T, Debeer A-, Kotaki Y, Rhodes L, Lundholm N, Sabbe K, Vyverman W (2010) Limits to gene flow in a cosmopolitan marine planktonic diatom. Proceedings of the National Academy of Sciences 107(29): 12952–12957. https://doi.org/10.1073/pnas.1001380107

Chambers J (2013) SoDA: Functions and Examples for "Software for Data Analysis". R package version 1.0-6. https://CRAN.R-project.org/package=SoDA [Accessed on: 2017-9-27]

Countway P, Gast R, Dennett M, Savai P, Rose J, Caron D (2007) Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). Environmental Microbiology 9(5): 1219–1232. https://doi.org/10.1111/j.1462-2920.2007.01243.x

Countway PD, Gast RJ, Savai P, Caron DA (2005) Protistan diversity estimates based on 18S rDNA from seawater incubations in the Western North Atlantic. Journal of Eukaryotic Microbiology 52(2): 95–106. https://doi.org/10.1111/j.1550-7408.2005.05202006.x

Degnan PH, Ochman H (2011) Illumina-based analysis of microbial community diversity. The ISME Journal 6(1): 183–194. https://doi.org/10.1038/ismej.2011.74

de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney C, Le Bescot N, Probert I, Carmichael M, Poulain J, Romac S, Colin S, Aury J, Bittner L, Chaffron S, Dunthorn M, Engelen S, Flegontova O, Guidi L, Horák A, Jaillon O, Lima-Mendez G, Lukeš J, Malviya S, Morard R, Mulot M, Scalco E, Siano R, Vincent F, Zingone A, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Acinas SG, Bork P, Bowler C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Not F, Ogata H, Pesant S, Raes J, Sieracki ME, Speich S, Stemmann L, Sunagawa S, Weissenbach J, Wincker P, Karsenti E (2015) Eukaryotic plankton diversity in the sunlit ocean. Science 348(6237): 1261605. https://doi.org/10.1126/science.1261605

Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software 22(4): 1–20. https://doi.org/10.18637/jss.v022.i04

Dray S, Blanchet G, Borcard D, Guenard G, Jombart T, Larocque G, Legendre P, Wagner H (2017) adespatial: Multivariate Multiscale Spatial Analysis. https://CRAN.R-project.org/package=adespatial [Accessed on: 2017-9-27]

Edgar R, Haas B, Clemente J, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27(16): 2194–2200. https://doi.org/10.1093/bioinformatics/btr381

Fernández L, Fournier B, Rivera R, Lara E, Mitchell ED, Hernández C (2016) Water-energy balance, past ecological perturbations and evolutionary constraints shape the latitudinal diversity gradient of soil testate amoebae in south-western South America. Global Ecology and Biogeography 25(10): 1216–1227. https://doi.org/10.1111/geb.12478

Fernández LD, Hernández CE, Schiaffino MR, Izaguirre I, Lara E (2017) Geographical distance and local environmental conditions drive the genetic population structure of a freshwater microalga (Bathycoccaceae; Chlorophyta) in Patagonian lakes. FEMS Microbiol Ecol 93. https://doi.org/doi.org/10.1093/femsec/fix125

Filker S, Sommaruga R, Vila I, Stoeck T (2016) Microbial eukaryote plankton communities of high-mountain lakes from three continents exhibit strong biogeographic patterns. Molecular Ecology 25(10): 2286–2301. https://doi.org/10.1111/mec.13633

Foissner W (1987) Soil protozoa: fundamental problems, ecological significance, adaptions in ciliates and testaceans, bioindicators and guide to the literature. Progress in Protistology 2: 69–212.

Foissner W (2006) Biogeography and dispersal of micro-organisms: a review emphasizing protists. Acta Protozoologica 45: 111–136.

Fonseca VG, Nichols B, Lallias D, Quince C, Carvalho GR, Power DM, Creer S (2012) Sample richness and genetic diversity as drivers of chimera formation in nSSU metagenetic analyses. Nucleic Acids Research 40 (9): e66-e66. https://doi.org/10.1093/nar/gks002

Gimmler A, Korn R, Vargas Cd, Audic S, Stoeck T (2016) The Tara Oceans voyage reveals global diversity and distribution patterns of marine planktonic ciliates. Scientific Reports 6: 33555. https://doi.org/10.1038/srep33555

Green J, Holmes A, Westoby M, Oliver I, Briscoe D, Dangerfield M, Gillings M, Beattie A (2004) Spatial scaling of microbial eukaryote diversity. Nature 432(7018): 747–750. https://doi.org/10.1038/nature03034

Grossmann L, Jensen M, Pandey R, Jost S, Bass D, Psenner R, Boenigk J (2016a) Molecular investigation of protistan diversity along an elevation transect of alpine lakes. Aquatic Microbial Ecology 78(1): 25–37. https://doi.org/10.1038/ismej.2016.10

Grossmann L, Jensen M, Heider D, Jost S, Glücksman E, Hartikainen H, Mahamdallie SS, Gardner M, Hoffmann D, Bass D, Boenigk J (2016b) Protistan community analysis: key findings of a large-scale molecular sampling. The ISME Journal 10(9): 2269–2279. https://doi.org/10.1038/ismej.2016.10.

Habbe KA (2007) Stratigraphical terms for the Quaternary of the south German Alpine Foreland. In: Litt T (Ed.) Stratigraphie von Deutschland – Quartär. Eiszeitalter und Gegenwart 56: 66–83.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405(6789): 907–913. https://doi.org/10.1038/35016000

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society B: Biological Sciences 359(1442): 183–195. https://doi.org/10.1098/rstb.2003.1388

Hillebrand H (2004) On the generality of the latitudinal diversity gradient. The American Naturalist 163(2): 192–211. https://doi.org/10.1086/381004

Huber J, Morrison H, Huse S, Neal P, Sogin M, Mark Welch D (2009) Effect of PCR amplicon size on assessments of clone library microbial diversity and community structure. Environmental Microbiology 11 (5): 1292–1302. https://doi.org/10.1111/j.1462-2920.2008.01857.x

Humboldt Av, Bonpland A (1807) Ideen zu einer Geographie der Pflanzen nebst einem Naturgemälde der Tropenländer, auf Beobachtungen und Messungen gegründet, welche vom 10ten Grade nördlicher bis zum 10ten Grade südlicher Breite, in den Jahren 1799, 1800, 1801, 1802 und 1803 angestellt worden sind /von Al. von Humboldt und A. Bonpland ; bearbeitet und herausgegeben von dem erstern. Cotta-Verlag, Tübingen. https://doi.org/10.5962/bhl.title.9310

Izaguirre I, Saad J, Schiaffino MR, Vinocur A, Tell G, Sánchez ML, Allende L, Sinistro R (2015) Drivers of phytoplankton diversity in Patagonian and Antarctic lakes across a latitudinal gradient (2150 km): the importance of spatial and environmental factors. Hydrobiologia 764(1): 157–170. https://doi.org/10.1007/s10750-015-2269-2

Jensen M (2017) V9_Clust.R, R-Script for modifying DNA-sequence-abundance tables: clustering of related sequences (e.g. SSU-ITS1) according to 100 % identical sub-sequences. https://github.com/manfred-uni-essen/V9-cluster [Accessed on: 2017-9-28]

Kammerlander B, Breiner H, Filker S, Sommaruga R, Sonntag B, Stoeck T (2015) High diversity of protistan plankton communities in remote high mountain lakes in the European Alps and the Himalayan mountains. FEMS Microbiology Ecology 91(4). https://doi.org/10.1093/femsec/fiv010

Katz L, McManus G, Snoeyenbos-West O, Griffin A, Pirog K, Costas B, Foissner W (2005) Reframing the 'Everything is everywhere' debate: evidence for high gene flow and diversity in ciliate morphospe-

cies. Aquatic Microbial Ecology 41: 55–65. https://doi.org/10.3354/ame041055

Lange A, Jost S, Heider D, Bock C, Budeus B, Schilling E, Strittmatter A, Boenigk J, Hoffmann D (2015) AmpliconDuo: a split-sample filtering protocol for high-throughput amplicon sequencing of microbial communities. PLoS ONE 10(11): e0141590. https://doi.org/10.1371/journal.pone.0141590

Lara E, Mitchell ED, Moreira D, García PL (2011) Highly diverse and seasonally dynamic protist community in a Pristine Peat Bog. Protist 162(1): 14–32. https://doi.org/10.1016/j.protis.2010.05.003

Legendre P (2014a) R package to compute space-constrained or time-constrained agglomerative clustering from a dissimilarity matrix computed from multivariate data. https://github.com/philippec/fonctions_R_git/tree/master/const.clust [Accessed on: 2017-9-28]

Legendre P (2014b) Interpreting the replacement and richness difference components of beta diversity. Global Ecology and Biogeography 23(11): 1324–1334. https://doi.org/10.1111/geb.12207

Lepère C, Domaizon I, Debroas D (2007) Community composition of lacustrine small eukaryotes in hyper-eutrophic conditions in relation to top-down and bottom-up factors. FEMS Microbiology Ecology 61(3): 483–495. https://doi.org/10.1111/j.1574-6941.2007.00359.x

Litt T, Behre KE, Meyer KD, Stephan HJ, Wansa S (2007) Stratigraphical terms for the Quaternary of the North German glaciation area. In: Litt T (Ed.) Stratigraphie von Deutschland – Quartär. Eiszeitalter und Gegenwart 56: 7–65.

Livermore JA, Jones SE (2015) Local-global overlap in diversity informs mechanisms of bacterial biogeography. The ISME Journal 9(11): 2413–2422. https://doi.org/10.1038/ismej.2015.51

Logares R, Boltovskoy A, Bensch S, Laybourn-Parry J, Rengefors K (2009) Genetic diversity patterns in five protist species occurring in lakes. Protist 160(2): 301–317. https://doi.org/10.1016/j.protis.2008.10.004

López-García P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature 409(6820): 603–607. https://doi.org/10.1038/35054537

Lovejoy C, Massana R, Pedros-Alio C (2006) Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent Seas. Applied and Environmental Microbiology 72(5): 3085–3095. https://doi.org/10.1128/aem.72.5.3085-3095.2006

Mahé F, Rognes T, Quince C, Vargas Cd, Dunthorn M (2014) Swarm: robust and fast clustering method for amplicon-based studies. PeerJ 2. https://doi.org/10.7717/peerj.593

Mann D, Vanormelingen P (2013) An inordinate fondness? The number, distributions, and origins of diatom species. Journal of Eukaryotic Microbiology 60(4): 414–420. https://doi.org/10.1111/jeu.12047

Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD (2012) PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics 13(1): 31. https://doi.org/10.1186/1471-2105-13-31

Massana R, Pedrós-Alió C (2008) Unveiling new microbial eukaryotes in the surface ocean. Current Opinion in Microbiology 11(3): 213–218. https://doi.org/10.1016/j.mib.2008.04.004

Massana R, Campo Jd, Sieracki ME, Audic S, Logares R (2013) Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. The ISME Journal 8(4): 854–866. https://doi.org/10.1038/ismej.2013.204

Massana R, Gobet A, Audic S, Bass D, Bittner L, Boutte C, Chambouvet A, Christen R, Claverie J, Decelle J, Dolan J, Dunthorn M, Edvardsen B, Forn I, Forster D, Guillou L, Jaillon O, F. Kooistra WC, Logares R, Mahé F, Not F, Ogata H, Pawlowski J, Pernice M, Probert I, Romac S, Richards T, Santini S, Shalchian-Tabrizi K, Siano R, Simon N, Stoeck T, Vaulot D, Zingone A, Vargas Cd (2015) Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. Environmental Microbiology 17(10): 4035–4049. https://doi.org/10.1111/1462-2920.12955

Moreira D, López-García P (2002) The molecular ecology of microbial eukaryotes unveils a hidden world. Trends in Microbiology 10(1): 31–38. https://doi.org/10.1016/s0966-842x(01)02257-0

Naselli-Flores L, Padisák J, Bach MF (2003) Phytoplankton and Equilibrium Concept: The Ecology of Steady-State Assemblages: Proceedings of the 13th Workshop of the International Association of Phytoplankton Taxonomy and Ecology (IAP). Springer, 403 pp.

Nogués-Bravo D, Araújo MB, Romdal T, Rahbek C (2008) Scale effects and human impact on the elevational species richness gradients. Nature 453(7192): 216–219. https://doi.org/10.1038/nature06812

Nolte V, Pandey RV, Jost S, Medinger R, Oottenwälder B, Boenigk J, Schlötterer C (2010) Contrasting seasonal niche separation between rare and abundant taxa conceals the extent of protist diversity. Molecular Ecology 19(14): 2908–2915. https://doi.org/10.1111/j.1365-294x.2010.04669.x

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D (2017) vegan: Community Ecology Package. CRAN. https://CRAN.R-project.org/package=vegan [Accessed on: 2017-9-28]

Pomeroy L (1974) The Ocean's Food Web, A Changing Paradigm. BioScience 24(9): 499–504. https://doi.org/10.2307/1296885

Renka RJ, Gebhardt A, Eglen S, Zuyev S, White D (2016) tripack: Triangulation of irregularly spaced data. R package version 1.3-8. https://CRAN.R-project.org/package=tripack [Accessed on: 2017-9-28]

Romdal T, Grytnes J (2007) An indirect area effect on elevational species richness patterns. Ecography 30(3): 440–448. https://doi.org/10.1111/j.0906-7590.2007.04954.x

Santucci F, Emerson BC, Hewitt GM (1998) Mitochondrial DNA phylogeography of European hedgehogs. Molecular Ecology 7(9): 1163–1172. https://doi.org/10.1046/j.1365-294x.1998.00436.x

Schiaffino MR, Lara E, Fernández L, Balagué V, Singer D, Seppey CW, Massana R, Izaguirre I (2016) Microbial eukaryote communities exhibit robust biogeographical patterns along a gradient of Patagonian and Antarctic lakes. Environmental Microbiology 18(12): 5249–5264. https://doi.org/10.1111/1462-2920.13566

Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27 (6): 863–864. https://doi.org/10.1093/bioinformatics/btr026

Schmitt T (2009) Biogeographical and evolutionary importance of the European high mountain systems. Frontiers in Zoology 6(1): 9. https://doi.org/10.1186/1742-9994-6-9

Scoble JM, Cavalier-Smith T (2014) Scale evolution in Paraphysomonadida (Chrysophyceae): sequence phylogeny and revised taxonomy of Paraphysomonas, new genus Clathromonas, and 25 new species. European Journal of Protistology 50(5): 551–592. https://doi.org/10.1016/j.ejop.2014.08.001

Sherr B, Sherr E, Caron D, Vaulot D, Worden A (2007) Oceanic protists. Oceanography 20(2): 130–134. https://doi.org/10.5670/oceanog.2007.57

Simon M, López-García P, Deschamps P, Moreira D, Restoux G, Bertolino P, Jardillier L (2015) Marked seasonality and high spatial variability of protist communities in shallow freshwater systems. The ISME Journal 9(9): 1941–1965. https://doi.org/10.1038/ismej.2015.6

Slapeta J, Moreira D, Lopez-Garcia P (2005) The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. Proceedings of the Royal Society B: Biological Sciences 272(1576): 2073–2081. https://doi.org/10.1098/rspb.2005.3195

Sommaruga R (2001) The role of solar UV radiation in the ecology of alpine lakes. Journal of Photochemistry and Photobiology B: Biology 62: 35–42. https://doi.org/10.1016/s1011-1344(01)00154-3

Sonntag B, Summerer M, Sommaruga R (2010) Factors involved in the distribution pattern of ciliates in the water column of a transparent alpine lake. Journal of Plankton Research 33(3): 541–546. https://doi.org/10.1093/plankt/fbq117

Sprengel K, de Candolle AP (1819) Grundzüge der wissenschaftlichen Pflanzenkunde: Zu Vorlesungen. C.Cnobloch, Leipzig.

Taib N, Mangot J, Domaizon I, Bronner G, Debroas D (2013) Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: new insights into the freshwater protist diversity. PLoS ONE 8(3): e58950. https://doi.org/10.1371/journal.pone.0058950

Tolotti M, Thies H, Cantonati M, Hansen CE, Thaler B (2003) Flagellate algae (Chrysophyceae, Dinophyceae, Cryptophyceae) in 48 high mountain lakes of the Northern and Southern slope of the Eastern Alps: biodiversity, taxa distribution and their driving variables. Hydrobiologia 502: 331–348. https://doi.org/10.1023/b:hydr.0000004291.03882.f7

Tolotti M, Manca M, Angeli N, Morabito G, Thaler B, Rott E, Stuchlik E (2006) Phytoplankton and Zooplankton Associations in a Set of Alpine High Altitude Lakes: Geographic Distribution and Ecology. Hydrobiologia 562(1): 99–122. https://doi.org/10.1007/s10750-005-1807-8

Triadó-Margarit X, Casamayor E (2012) Genetic diversity of planktonic eukaryotes in high mountain lakes (Central Pyrenees, Spain). Environmental Microbiology 14(9): 2445–2456. https://doi.org/10.1111/j.1462-2920.2012.02797.x

van der Gast C (2014) Microbial biogeography: the end of the ubiquitous dispersal hypothesis? Environmental Microbiology 17(3): 544–546. https://doi.org/10.1111/1462-2920.12635

Wit Rd, Bouvier T (2006) 'Everything is everywhere, but, the environment selects'; what did Baas Becking and Beijerinck really say? Environmental Microbiology 8(4): 755–758. https://doi.org/10.1111/j.1462-2920.2006.01017.x

Worden A, Cuvelier M, Bartlett D (2006) In-depth analyses of marine microbial community genomics. Trends in Microbiology 14(8): 331–336. https://doi.org/10.1016/j.tim.2006.06.008

Wu Q, Chatzinotas A, Wang J, Boenigk J (2009) Genetic diversity of eukaryotic plankton assemblages in eastern Tibetan lakes differing by their salinity and altitude. Microbial Ecology 58(3): 569–581. https://doi.org/10.1007/s00248-009-9526-8

Supplementary materials

Supplementary material 1
Sample Identifier for molecular analyses
Authors: Jens Boenigk, Sabina Wodniok, Christina Bock, Daniela Beisser, Christopher Hempel, Lars Grossmann, Anja Lange, Manfred Jensen
Data type: molecular analyses
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/10.3897/mbmg.2.21519.suppl1

Supplementary material 2
Lake characteristics and diversity indices
Authors: Jens Boenigk, Sabina Wodniok, Christina Bock, Daniela Beisser, Christopher Hempel, Lars Grossmann, Anja Lange, Manfred Jensen
Data type: lake characteristics and diversity indices
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/10.3897/mbmg.2.21519.suppl2

Supplementary material 3
Bray distance matrix
Authors: Jens Boenigk, Sabina Wodniok, Christina Bock, Daniela Beisser, Christopher Hempel, Lars Grossmann, Anja Lange, Manfred Jensen
Data type: Bray distance matrix
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/10.3897/mbmg.2.21519.suppl3

Supplementary material 4
Richness at different elevations
Authors: Jens Boenigk, Sabina Wodniok, Christina Bock, Daniela Beisser, Christopher Hempel, Lars Grossmann, Anja Lange, Manfred Jensen
Data type: Box plots
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/10.3897/mbmg.2.21519.suppl4